

0570

0731

A8



ENTERED

OIPE

## RAW SEQUENCE LISTING

DATE: 08/05/2002

PATENT APPLICATION: US/09/891,138A

TIME: 15:48:54

Input Set : N:\Crif3\07252002\I891138.raw

Output Set: N:\CRF4\08052002\I891138A.raw

```

1 <110> APPLICANT: Lin, Daniel Chi-Hong
2      Zhao, Jiagang
3      Chen, Jin-Long
4      Cutler, Gene
5      Tularik Inc.
6 <120> TITLE OF INVENTION: Novel Receptors
7 <130> FILE REFERENCE: 018781-006210US
8 <140> CURRENT APPLICATION NUMBER: US/09/891,138A
9 <141> CURRENT FILING DATE: 2001-06-25
10 <150> PRIOR APPLICATION NUMBER: US 60/213,461
11 <151> PRIOR FILING DATE: 2000-06-23
12 <160> NUMBER OF SEQ ID NOS: 26
13 <170> SOFTWARE: PatentIn Ver. 2.1
15 <210> SEQ ID NO: 1
16 <211> LENGTH: 1543
17 <212> TYPE: DNA
18 <213> ORGANISM: Mus musculus
19 <220> FEATURE:
20 <221> NAME/KEY: CDS
21 <222> LOCATION: (44)..(997)
22 <223> OTHER INFORMATION: mouse TGR18 G-protein coupled receptor (GPCR)
23 <400> SEQUENCE: 1
24      gctcctggca gagttttctg tcgagacaga agccgacagc agaatggcac agaattttatc 60
25      ttgtgagaat tggttggcaa cagaggctat ctggaataag tactacctct ctgcatttta 120
26      tgcaatcgag ttcatttttg gactgcttgg gaatgtcact gtggtgttcg gctacctctt 180
27      ctgcatgaag aactggaaca gcagcaatgt ctatcttttt aacctttcca tctctgactt 240
28      tgctttcctg tgcacccttc ccatcctgat aaagagttat gccaatgata aggggacctt 300
29      tggagatgtt ctctgtataa gcaaccgata tgtgcttcac accaacctct acaccagcat 360
30      cctcttcctc actttcatta gcatggaccg atatctgctc atgaagtacc ctttccgaga 420
31      acactttcta caaaagaagg aatttgccat tttaatctcg ctggctgtct gggccttagt 480
32      gaccttagaa gttctaccca tgctcacttt catcaattct gtcccaaaag aagagggcag 540
33      taactgcac gactatgcaa gttctggaac cctgaacac aatctcattt acagcctctg 600
34      cctgactttg ttgggcttcc taattcctct ctctgtgatg tgcttcttct actacaagat 660
35      ggtagtcttc ttaaagagga ggagccagca gcaagcaact gccttgccac tggacaaacc 720
36      ccaacgcctg gtggtcctgg cgggtgtgat cttctctata ctcttcacac cctatcatat 780
37      catgcgcaat ttgaggatcg cctcacgcct ggatagtgtg ccacaaggat gtacacagaa 840
38      ggccatcaaa tctatataca cactgacacg gcctctggcc tttctgaaca gtgccatcaa 900
39      tcccatcttc tacttcctca tgggagacca ttacagagag atgctgatta gtaagttcag 960
40      acaatacttc aagtccctta catccttcag gacatgagct gctggatgca ggtcttcaact 1020
41      cagccaaaat gagacacttg ataaacagtg ctgtgcagtt gagttttaac taagtaaacc 1080
42      accatttcta ggcttttagct ttccaccatc ctccaacccc cagggctgga gtacaagctg 1140
43      ggtccacatg aatcagaagg cagctctctg ttctgatttt aggttatacc cagagtatgg 1200
44      aaaaaataag gcatgagaaa gcattgacat cttcacttaa gaactgaaca aaagagaaca 1260

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/891,138A

DATE: 08/05/2002

TIME: 15:48:54

Input Set : N:\Crf3\07252002\I891138.raw

Output Set: N:\CRF4\08052002\I891138A.raw

```

45  aatattgtca atgtttggac acttaggatc tgaaatcttg gaaattttta gacctctttt 1320
46  tctatcagtg taaaaggaat acaagatagc tagttgcaaa tgctgaatgc atttcatcat 1380
47  tggtcaggtc gataagcgtg tttctgaaat agtcttattt ttattcttgt aatattaaaa 1440
48  tttatgtgaa aaatgaatat aattcaatgt acaacattag attttctatt tgaaaattat 1500
49  atttcttgaa aaaataactg ctgtgcctaa ataatcaat ata 1543
51 <210> SEQ ID NO: 2
52 <211> LENGTH: 317
53 <212> TYPE: PRT
54 <213> ORGANISM: Mus musculus
55 <220> FEATURE:
56 <223> OTHER INFORMATION: mouse TGR18 G-protein coupled receptor (GPCR)
57 <400> SEQUENCE: 2
58  Met Ala Gln Asn Leu Ser Cys Glu Asn Trp Leu Ala Thr Glu Ala Ile
59      1          5          10          15
60  Leu Asn Lys Tyr Tyr Leu Ser Ala Phe Tyr Ala Ile Glu Phe Ile Phe
61      20          25          30
62  Gly Leu Leu Gly Asn Val Thr Val Val Phe Gly Tyr Leu Phe Cys Met
63      35          40          45
64  Lys Asn Trp Asn Ser Ser Asn Val Tyr Leu Phe Asn Leu Ser Ile Ser
65      50          55          60
66  Asp Phe Ala Phe Leu Cys Thr Leu Pro Ile Leu Ile Lys Ser Tyr Ala
67      65          70          75          80
68  Asn Asp Lys Gly Thr Tyr Gly Asp Val Leu Cys Ile Ser Asn Arg Tyr
69      85          90          95
70  Val Leu His Thr Asn Leu Tyr Thr Ser Ile Leu Phe Leu Thr Phe Ile
71      100         105         110
72  Ser Met Asp Arg Tyr Leu Leu Met Lys Tyr Pro Phe Arg Glu His Phe
73      115         120         125
74  Leu Gln Lys Lys Glu Phe Ala Ile Leu Ile Ser Leu Ala Val Trp Ala
75      130         135         140
76  Leu Val Thr Leu Glu Val Leu Pro Met Leu Thr Phe Ile Asn Ser Val
77      145         150         155         160
78  Pro Lys Glu Glu Gly Ser Asn Cys Ile Asp Tyr Ala Ser Ser Gly Asn
79      165         170         175
80  Pro Glu His Asn Leu Ile Tyr Ser Leu Cys Leu Thr Leu Leu Gly Phe
81      180         185         190
82  Leu Ile Pro Leu Ser Val Met Cys Phe Phe Tyr Tyr Lys Met Val Val
83      195         200         205
84  Phe Leu Lys Arg Arg Ser Gln Gln Gln Ala Thr Ala Leu Pro Leu Asp
85      210         215         220
86  Lys Pro Gln Arg Leu Val Val Leu Ala Val Val Ile Phe Ser Ile Leu
87      225         230         235         240
88  Phe Thr Pro Tyr His Ile Met Arg Asn Leu Arg Ile Ala Ser Arg Leu
89      245         250         255
90  Asp Ser Trp Pro Gln Gly Cys Thr Gln Lys Ala Ile Lys Ser Ile Tyr
91      260         265         270
92  Thr Leu Thr Arg Pro Leu Ala Phe Leu Asn Ser Ala Ile Asn Pro Ile
93      275         280         285
94  Phe Tyr Phe Leu Met Gly Asp His Tyr Arg Glu Met Leu Ile Ser Lys

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/891,138A

DATE: 08/05/2002

TIME: 15:48:54

Input Set : N:\Crif3\07252002\I891138.raw

Output Set: N:\CRF4\08052002\I891138A.raw

```

95          290          295          300
96      Phe Arg Gln Tyr Phe Lys Ser Leu Thr Ser Phe Arg Thr
97      305          310          315
99 <210> SEQ ID NO: 3
100 <211> LENGTH: 1305
101 <212> TYPE: DNA
102 <213> ORGANISM: Homo sapiens
103 <220> FEATURE:
104 <221> NAME/KEY: CDS
105 <222> LOCATION: (1)..(1305)
106 <223> OTHER INFORMATION: human TGR21 G-protein coupled receptor (GPCR)
107 <400> SEQUENCE: 3
108      atggaggatc tctttagccc ctcaattctg ccgcccggcgc ccaacatttc cgtgcccata 60
109      ttgctgggct ggggtctcaa cctgaccttg gggcaaggag cccctgcctc tgggccgccc 120
110      agccgcccgc tccgcctggt gttcctgggg gtcactctgg tggtagcggt ggcaggcaac 180
111      accacagtgc tgtgccgcct gtgcggcggc ggccggccct gggcgggccc caagcgtcgc 240
112      aagatggact tcctgctggt gcagctggcc ctggcggacc tgtacgcgtg cgggggcacg 300
113      gcgctgtcac agctggcctg ggaactgctg ggcgagcccc gcgcggccac gggggacctg 360
114      gcgtgccgct tcctgcagct gctgcaggca tccgggcggg gcgcctcggc ccacctcgtg 420
115      gtgctcatcg cctcgcagcg ccggcgcgcg gtgcgtcttc cgcacggccg gccgctgccc 480
116      gcgcgtgccc tcgcgcctt gggtggctg ctggcaactgc tgcaggcgct gcccccgccc 540
117      ttcgtggtgc gcggggactc cccctgcgcg ctgccgcgcg cgcgcgcgcg aacgtccctg 600
118      cagccaggcg cgccccggc cgcccgcgcc tggccggggg agcgtcgtg ccacgggata 660
119      ttcgcgcccc tgccgcgtg gcacctgcag gtctacgcgt tctacgaggc cgtcgcgggc 720
120      ttcgtcgcgc ctgttacggt cctgggcgtc gcttgcggcc acctactctc cgtctggtgg 780
121      cggcacccgc cgcaggcccc cgcggtgca gcgcctggt cggcgagccc aggtcgagcc 840
122      cctgcgcccc gcgcgtgcc ccgcgccaag gtgcagagcc tgaagatgag cctgctgctg 900
123      gcgctgctgt tcgtgggctg cgagctgccc tactttgccg cccggctggc gcccgctgg 960
124      tcgtccgggc ccgcgggaga ctgggaggga gagggcctgt cggcggcgct gcgcgtggtg 1020
125      gcgatggcca acagcgtct caatcccttc gtctacctct tcttcaggc gggcgactgc 1080
126      cggctccggc gacagctgc gaagcggctg ggctctctgt gctgcgcgcc gcagggaggc 1140
127      gcggaggacg aggaggggcc ccggggccac caggcgctct accgccaacg ctggccccac 1200
128      cctcattatc accatgctcg gcgggaaccg ctggacgagg gcggcttgcg cccacccctt 1260
129      ccgcgcccc gaccctgcc ttgctcctgc gaaagtgcct tctag 1305
131 <210> SEQ ID NO: 4
132 <211> LENGTH: 434
133 <212> TYPE: PRT
134 <213> ORGANISM: Homo sapiens
135 <220> FEATURE:
136 <223> OTHER INFORMATION: human TGR21 G-protein coupled receptor (GPCR)
137 <400> SEQUENCE: 4
138      Met Glu Asp Leu Phe Ser Pro Ser Ile Leu Pro Pro Ala Pro Asn Ile
139      1          5          10          15
140      Ser Val Pro Ile Leu Leu Gly Trp Gly Leu Asn Leu Thr Leu Gly Gln
141      20          25          30
142      Gly Ala Pro Ala Ser Gly Pro Pro Ser Arg Arg Val Arg Leu Val Phe
143      35          40          45
144      Leu Gly Val Ile Leu Val Val Ala Val Ala Gly Asn Thr Thr Val Leu
145      50          55          60

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/891,138A

DATE: 08/05/2002

TIME: 15:48:54

Input Set : N:\Crif3\07252002\I891138.raw

Output Set: N:\CRF4\08052002\I891138A.raw

```

146 Cys Arg Leu Cys Gly Gly Gly Gly Pro Trp Ala Gly Pro Lys Arg Arg
147 65 70 75 80
148 Lys Met Asp Phe Leu Leu Val Gln Leu Ala Leu Ala Asp Leu Tyr Ala
149 85 90 95
150 Cys Gly Gly Thr Ala Leu Ser Gln Leu Ala Trp Glu Leu Leu Gly Glu
151 100 105 110
152 Pro Arg Ala Ala Thr Gly Asp Leu Ala Cys Arg Phe Leu Gln Leu Leu
153 115 120 125
154 Gln Ala Ser Gly Arg Gly Ala Ser Ala His Leu Val Val Leu Ile Ala
155 130 135 140
156 Leu Glu Arg Arg Arg Ala Val Arg Leu Pro His Gly Arg Pro Leu Pro
157 145 150 155 160
158 Ala Arg Ala Leu Ala Ala Leu Gly Trp Leu Leu Ala Leu Leu Leu Ala
159 165 170 175
160 Leu Pro Pro Ala Phe Val Val Arg Gly Asp Ser Pro Ser Pro Leu Pro
161 180 185 190
162 Pro Pro Pro Pro Thr Ser Leu Gln Pro Gly Ala Pro Pro Ala Ala
163 195 200 205
164 Arg Ala Trp Pro Gly Glu Arg Arg Cys His Gly Ile Phe Ala Pro Leu
165 210 215 220
166 Pro Arg Trp His Leu Gln Val Tyr Ala Phe Tyr Glu Ala Val Ala Gly
167 225 230 235 240
168 Phe Val Ala Pro Val Thr Val Leu Gly Val Ala Cys Gly His Leu Leu
169 245 250 255
170 Ser Val Trp Trp Arg His Arg Pro Gln Ala Pro Ala Ala Ala Pro
171 260 265 270
172 Trp Ser Ala Ser Pro Gly Arg Ala Pro Ala Pro Ser Ala Leu Pro Arg
173 275 280 285
174 Ala Lys Val Gln Ser Leu Lys Met Ser Leu Leu Leu Ala Leu Leu Phe
175 290 295 300
176 Val Gly Cys Glu Leu Pro Tyr Phe Ala Ala Arg Leu Ala Ala Ala Trp
177 305 310 315 320
178 Ser Ser Gly Pro Ala Gly Asp Trp Glu Gly Glu Gly Leu Ser Ala Ala
179 325 330 335
180 Leu Arg Val Val Ala Met Ala Asn Ser Ala Leu Asn Pro Phe Val Tyr
181 340 345 350
182 Leu Phe Phe Gln Ala Gly Asp Cys Arg Leu Arg Arg Gln Leu Arg Lys
183 355 360 365
184 Arg Leu Gly Ser Leu Cys Cys Ala Pro Gln Gly Gly Ala Glu Asp Glu
185 370 375 380
186 Glu Gly Pro Arg Gly His Gln Ala Leu Tyr Arg Gln Arg Trp Pro His
187 385 390 395 400
188 Pro His Tyr His His Ala Arg Arg Glu Pro Leu Asp Glu Gly Gly Leu
189 405 410 415
190 Arg Pro Pro Pro Pro Arg Pro Arg Pro Leu Pro Cys Ser Cys Glu Ser
191 420 425 430
192 Ala Phe
194 <210> SEQ ID NO: 5
195 <211> LENGTH: 1266

```

## RAW SEQUENCE LISTING

DATE: 08/05/2002

PATENT APPLICATION: US/09/891,138A

TIME: 15:48:54

Input Set : N:\Crif3\07252002\I891138.raw

Output Set: N:\CRF4\08052002\I891138A.raw

```

196 <212> TYPE: DNA
197 <213> ORGANISM: Homo sapiens
198 <220> FEATURE:
199 <221> NAME/KEY: CDS
200 <222> LOCATION: (25)..(1197)
201 <223> OTHER INFORMATION: human TGR62 G-protein coupled receptor (GPCR)
202 <400> SEQUENCE: 5
203      tgaccttctt catcatttga tgtgatgcc aataactaata gcacaatcaa tttatcacta 60
204      agcactcgtg ttacttttagc attttttatg tccttagtag cttttgctat aatgctagga 120
205      aatgctttgg tcatttttagc ttttgtggtg gacaaaaacc ttagacatcg aagtagttat 180
206      ttttttctta acttggccat ctctgacttc tttgtgggtg tgatctccat tcctttgtac 240
207      atccctcaca cgctgttcga atgggatttt ggaaaggaaa tctgtgtatt ttggctcact 300
208      actgactatc tgttatgtac agcatctgta tataacattg tcctcatcag ctatgatcga 360
209      tacctgtcag tctcaaatgc tgtgtcctat agaactcaac atactggggg cttgaagatt 420
210      gttactctga tgggtggcgt ttgggtgctg gccttcttag tgaatgggcc aatgattcta 480
211      gtttcagagt cttggaagga tgaaggtagt gaatgtgaac ctggattttt ttcggaatgg 540
212      tacatccttg ccatcacatc attccttgga ttcgtgatcc cagtcattct agtcgcttat 600
213      ttcaacatga atatttattg gagcctgtgg aagcgtgatc atctcagtag gtgccaaagc 660
214      catcctggac tgactgctgt ctcttccaac atctgtggac actcattcag aggtagacta 720
215      tcttcaagga gatctctttc tgcacgcaca gaagttcctg catcctttca ttcagagaga 780
216      cagaggagaa agagtgtctc catgttttcc tcaagaacca agatgaatga caatacaatt 840
217      gcttccaaaa tgggttcctt ctcccaatca gattctgtag ctcttcacca aagggaacat 900
218      gttgaactgc ttagagccag gagattagcc aagtcactgg ccattctctt aggggttttt 960
219      gctgtttgct gggctccata ttctctgttc acaattgtcc tttcatttta ttcctcagca 1020
220      acaggctcta aatcagtttg gtatagaatt gcattttggc ttcagtgggt caattccttt 1080
221      gtcaatcctc ttttgtatcc attgtgtcac aagcgctttc aaaaggcttt cttgaaaata 1140
222      ttttgtataa aaaagcaacc tctaccatca caacacagtc ggtcagtatc ttcttaaaga 1200
223      caattttctc acctctgtaa attttagtct caatctcacc taaatgaatc aggtctgccc 1260
224      tttatc
226 <210> SEQ ID NO: 6
227 <211> LENGTH: 390
228 <212> TYPE: PRT
229 <213> ORGANISM: Homo sapiens
230 <220> FEATURE:
231 <223> OTHER INFORMATION: human TGR62 G-protein coupled receptor (GPCR)
232 <400> SEQUENCE: 6
233      Met Pro Asp Thr Asn Ser Thr Ile Asn Leu Ser Leu Ser Thr Arg Val
234      1          5          10          15
235      Thr Leu Ala Phe Phe Met Ser Leu Val Ala Phe Ala Ile Met Leu Gly
236      20          25          30
237      Asn Ala Leu Val Ile Leu Ala Phe Val Val Asp Lys Asn Leu Arg His
238      35          40          45
239      Arg Ser Ser Tyr Phe Phe Leu Asn Leu Ala Ile Ser Asp Phe Phe Val
240      50          55          60
241      Gly Val Ile Ser Ile Pro Leu Tyr Ile Pro His Thr Leu Phe Glu Trp
242      65          70          75          80
243      Asp Phe Gly Lys Glu Ile Cys Val Phe Trp Leu Thr Thr Asp Tyr Leu
244      85          90          95
245      Leu Cys Thr Ala Ser Val Tyr Asn Ile Val Leu Ile Ser Tyr Asp Arg

```

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/09/891,138A

DATE: 08/05/2002  
TIME: 15:48:55

Input Set : N:\Crf3\07252002\I891138.raw  
Output Set: N:\CRF4\08052002\I891138A.raw

Invalid Line Length:

The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:13; Line(s) 572

Seq#:14; Line(s) 600

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/891,138A

DATE: 08/05/2002

TIME: 15:48:55

Input Set : N:\Crf3\07252002\I891138.raw

Output Set: N:\CRF4\08052002\I891138A.raw